A Framework for Teaching Impedance Behaviours by Combining Human and Robot 'Best Practice'

Yuchen Zhao, Aran Sena, Fan Wu and Matthew J. Howard[1*]

***Abstract*— This paper presents a programming by demonstration framework for teaching impedance modulation using human demonstrations. Physiologically, human stiffness and damping are coupled at the muscle level, restricting the ability to modulate impedance according to task demands. Robotic systems often do not have this restriction (stiffness and damping can be varied independently), but the challenge is to devise an appropriate variable impedance profile for a given task. In this paper, the task critical component is first learned for imitation and a robot-specific controller is then blended into the control using the null space. In doing so, the control cheme takes advantage of both human and robot 'best practice'. Experimental results on a physical robot suggest an order of magnitude better mean performance, with lower variance, can be achieved using the blended scheme.**

I. Introduction

By virtue of a combination of passive dynamic properties of muscles and tendons and the redundant actuation of the antagonistic musculoskeletal system, humans are prime examples of systems able to control not only the kinematics of movement, but also force and impedance. They do this both through muscular co-contraction and posture selection [1]. In an attempt to reproduce this versatility, much recent emphasis has been placed on the development of so-called *soft robots*, namely, those that incorporate compliance in design, either physically or through fast force feedback control.

However, improving the design of robotic hardware is only one part of puzzle: human versatility is also due, in large part, to *sophisticated control strategies* that must be somehow reproduced in these new robotic systems. The reproduction of human impedance modulation skills is attractive, but human behaviour is highly optimised to the specific properties of the body. While this makes it highly robust and versatile, imitating all aspects of their movement can be *disadvantageous* for robotic systems whose embodiments have different dynamic properties.

From the biological literature, it is known that the stiffness and damping are *physiologically coupled* in human and animal behaviour [2]: damping increases in proportion to stiffness. In many cases, this coupling is beneficial, for example, to maintain stability. However, in other cases, the benefit is less clear. For example, in dynamic or explosive movements, such as running, throwing or hitting, it is known that introducing spring-like variable impedance actuation enables energy storage and can enhance throwing ability [3].

[1]Yuchen Zhao, Aran Sena, Fan Wu and Matthew J. Howard are with the Center for Robotics Research, Department of Informatics, King's College London, London, UK. matthew.j.howard@kcl.ac.uk.

*This work is supported by the Engineering and Physical Sciences Research Council (EPSRC) SoftSkills project, EP/P010202/1.

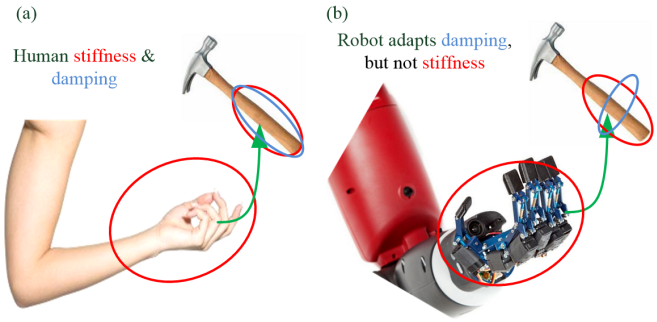

Fig. 1: Augmenting human impedance modulation in handle grasping: (a) Human impedance modulation is coupled, resulting in an alignment of stiffness and damping ellipses system may cause dissipation of energy. (b) Robotic systems can independently vary damping while still following the desired human stiffness profile.

In humans, however, peaks in stiffness profiles (associated with enhanced energy storage) must also result in peaks in damping (associated with *energy dissipation*), negating (or at least reducing) the benefits of storage.

Such examples show that only imitating parts of the human demonstrated behaviour (*i.e.,* those related to the task goals), while replacing the redundant parts with robot-specific controllers, may yield better task performance. Therefore, in the impedance modulation tasks such as throwing and reaching (as shown in Fig.1), it is desirable the robot imitates the stiffness from human but adapts the damping strategy other than human.

This motivates the idea to combine human and robot impedance modulations in a data-driven fashion, automatically extracting task redundancies and replacing idiosyncratic human behaviours with controllers optimised for the robot.

In this paper, a methodology is proposed to fulfil these motivations by (i) gathering human impedance demonstrations, (ii) performing a task redundancies analysis, and (iii) blending the demonstrated behaviour with a optimal control scheme for execution by the robot. Its contribution is to demonstrate that task relevant human impedance modulation strategy can be learnt through *imitation*, while that for secondary task is *decoupled*, and tailored instead to the robot's dynamics. In doing so, it establishes a way to blend both human and robot '*best practice*' in impedance control.

II. Problem Definition

As introduced above, imitating the impedance modulation behaviour of the human is attractive but may be suboptimal for the robot due to the differences in embodiment. In order to take advantage of both human and robot 'best practice', it is desirable to decompose the task related component

from the impedance observations and optimise only in the secondary task by using robot controller. In this section, the operational space impedance controller is formulated and a line-reaching task is illustrated as an example.

A. Operational Space Impedance Control

In impedance control, the desired operational space controller is represented by:

$$\mathbf{f}_s = \mathbf{K}_s(\mathbf{x}_d - \mathbf{x}) + \mathbf{D}_s(\dot{\mathbf{x}}_d - \dot{\mathbf{x}}) \tag{1}$$

where $\mathbf{x}, \mathbf{x}_d \in \mathbb{R}^{\mathcal{M}}$ are the position and desired position in the operational space ($\mathcal{M}$ is the dimension of the operational space). $\mathbf{K}_s \in \mathbb{R}^{\mathcal{M}\times\mathcal{M}}$ is the desired stiffness, $\mathbf{D}_s \in \mathbb{R}^{\mathcal{M}\times\mathcal{M}}$ is the desired damping in operational space and $\mathbf{f}_s \in \mathbb{R}^{\mathcal{M}}$ is the restoring force vector.

The human impedance control dynamically regulates the relation between the restoring force $\mathbf{f}_c$ and state vector $\mathbf{z} = [\mathbf{x}^T, \dot{\mathbf{x}}^T]^T$ using impedance variables $\mathbf{K}_s$, $\mathbf{D}_s$ and equilibrium states $\mathbf{z}_d$.

In this paper, the control input is defined as $\mathbf{u} = [\mathbf{k}^T, \mathbf{d}^T]^T \in \mathbb{R}^{2\mathcal{M}}$. To take into account the coupling effect of the stiffness and damping, the diagonal entries of stiffness $\mathbf{K}$ and damping matrix $\mathbf{D}$ is denoted as $\mathbf{k} = [k_1, k_2, ..., k_{\mathcal{M}}]^T$ and $\mathbf{d} = [d_1, d_2, ..., d_{\mathcal{M}}]^T$. The coupling relation is defined in function:

$$\mathbf{d} = \mathbf{C}(\mathbf{k}) \tag{2}$$

B. Representation of the constraint system

The general representation of the constraint system is defined as following:

$$\mathbf{A}(\mathbf{k})\mathbf{y} = \mathbf{b}(\mathbf{k}) \tag{3}$$

where $\mathbf{A} \in \mathbb{R}^{\mathcal{S}\times\mathcal{P}}$ is the *constraint matrix* which projects the task space policy onto the relevant part of the control space. Inverting (3), results in the relation:

$$\mathbf{y} = \mathbf{A}^{\dagger}(\mathbf{k})\mathbf{b}(\mathbf{k}) + \mathbf{N}\boldsymbol{\pi}(\mathbf{k}) \tag{4}$$

where $\mathbf{A}^{\dagger} = \mathbf{A}^T(\mathbf{A}^T\mathbf{A})^{-1}$ denotes the unique Moore-Penrose pseudo-inverse of the matrix $\mathbf{A}$, $\mathbf{N} = \mathbf{I} - \mathbf{A}^{\dagger}\mathbf{A}$ ($\mathbf{I} \in R^{\mathcal{P}\times\mathcal{P}}$), and $\boldsymbol{\pi}(\mathbf{k}) \in \mathbb{R}^{\mathcal{P}}$ is the lower priority control policy working in the null-space.

C. A grasping example

In this section, a 2-dimensional handle grasping task is used to illustrate impedance modulation problem defined in this paper. As shown in Fig.1, the human demonstrator needs to grasp a handle in the operation space. The primary task demonstrated is maintaining the shape of stiffness (red ellipse) from human demonstration (left plot) while the secondary task is minimising energy along the handle axis.

Since the existence of the task redundancy along the handle axis (more disturbance is allowed in this direction) the impedance variables $\mathbf{u}$ can be represented with priorities (The representation of the constrained system is thoroughly studied and can be found in [4]). Also due to the coupling effect of the stiffness and damping, the task priority can be represented only using stiffness variable ($\mathbf{k}$). Therefore, the observation $\mathbf{y} \equiv \dot{\mathbf{k}}$. The state in task constraint $\mathbf{A}(\mathbf{k})$ is $\mathbf{k}$. The observation pairs for decomposing task component ($\dot{\mathbf{k}}_{ts}$) and null-space component ($\dot{\mathbf{k}}_{ns}$) of the stiffness variables are $[\mathbf{k}, \dot{\mathbf{k}}]^T$.

The high-priority task component ($\dot{\mathbf{k}}_{ts}$) in this handle grasping task are learned from human demonstrations and can be kept for robot imitation. The lower priority null-space component is suboptimal and redundant (*i.e.,* assuming a torque controlled robot with active impedance controller) due to the coupling effect ($\dot{\mathbf{k}}_{ns}$). It can be switched out and replaced by robot-specific controller. Therefore, the 'best practice' from both human and robot is combined. Details implementation of the methods is explained in section IV.

III. Background and Related work

Many PbD methods exist for learning and generalising tasks from data [5], including trajectory-based learning techniques (*i.e.,* hidden Markov models [6], Gaussian mixture models[7]), policy methods (*i.e.,* dynamic movement primitives[8]) and forward/inverse reinforcement learning [9]. Most methods for PbD have been developed primarily in the context of learning tasks in the kinematic domain. However, there has recently been increasing interest in applying these to soft robots for force and impedance control. For example, Kronander & Billard [10] present methods for haptic teaching of stiffness by applying perturbations during kinesthetic demonstrations. Khansari-Zadeh & Khatib [11] present a method for learning potential function representations of impedance behaviours, whereby "stiffness targets" are used with a regression approach. Petric et al. [12] apply dynamic movement primitives to form a database of compliant behaviours learnt from torque and kinematic data. Mori, MH, & Vijayakumar [13] apply inverse reinforcement learning to learn an objective function representation of a dynamic hitting task, using a comobination of Electromyography (EMG) and kinematic data. Similar data was used by Ajoudani, Tsagarakis, & Bicchi for estimating and tracking impedance behaviour in a teleoperation framework [14]. Rozo et al[15] has implemented a impedance teaching strategy in human robot cooperative task under position and force constraints.

With the exception of the latter three, all of these methods rely on the demonstrator's ability to select appropriate impedance control strategies from their own experience. However, this makes a strong assumption of users' knowledge that may result in inappropriate selection of strategies. This may cause the robot to work sub-optimally or worse, risk instabilities and danger for both the operator and the robot itself [16]. For the methods relying on the EMG data, the transfer of control strategies is implicit (since the user's impedance is reflected in the data), but then the issues of sub-optimality due to mismatched embodiment returns.

IV. Method

The methodology contains three parts: i) gathering human impedance demonstrations ii) task redundancies analysis of impedance and iii) blending new controller using optimal control scheme. Figure 2 shows the proposed learning and control scheme. User kinematics and impedance profiles will be recorded during demonstrations of a task. Then task will

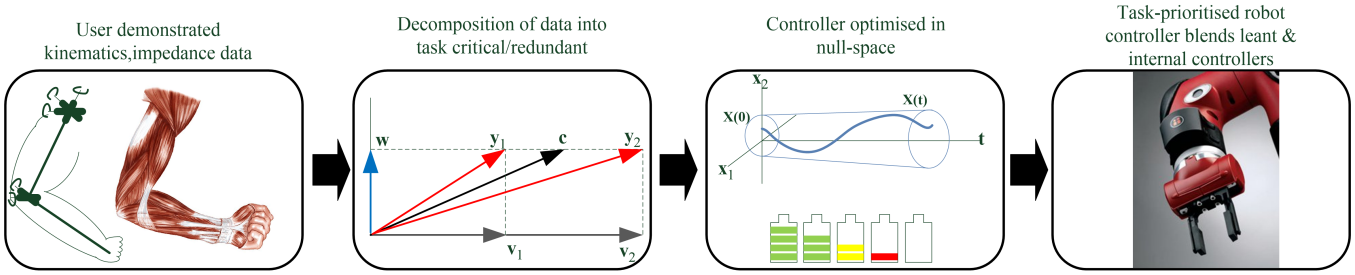

Fig. 2: The methodology of learning impedance modulations.

be decomposed into task critical and redundant components. Replacing the latter the robot-specific controllers will optimise over secondary task objective (*i.e.,* maximising stability and minimising energy). It will be blended in support of the task component.

A. Gathering human impedance demonstrations

In this paper, the human stiffness modulation in the task are derived from variations in the demonstrated trajectories [1]. More specifically, the position data of multiple motion trajectories are recorded and used to: (i) calculate the data covariance $\mathbf{\Sigma}$ using Gaussian mixture regression (GMR) and (ii) estimate stiffness $\mathbf{K}$ using data covariance $\mathbf{\Sigma}$.

In the first step, the data covariance matrix $\mathbf{\Sigma}$ can be estimated using the Gaussian mixture model (GMM) and GMR method [18]. The learned GMM is time-dependent and outputs μ_t and covariance $\mathbf{\Sigma}_t$ of position $\mathbf{x}$ at time t.

Given the covariance matrix $\mathbf{\Sigma}_t$ estimated, the stiffness matrix $\mathbf{K}_t$ (at time t) can be estimated by first decomposing the covariance matrix $\mathbf{\Sigma}_t$ as eigen vector matrix $\mathbf{V}$ and eigen value matrix $\mathbf{\Lambda}$ as follow (for simplicity, the subscripts t are dropped):

$$\mathbf{\Sigma} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T \in \mathbb{R}^{\mathcal{S}\times\mathcal{S}} \tag{5}$$

and then compose $\mathbf{K}$ in operational space:

$$\mathbf{K} = \mathbf{V}\mathbf{\Gamma}\mathbf{V}^T \in \mathbb{R}^{\mathcal{S}\times\mathcal{S}} \tag{6}$$

where the diagonal entries of the $\mathbf{\Gamma}$ matrix is inversely proportional to the square root of the corresponding eigenvalue (*i.e.,* $\sigma = \sqrt{\lambda}$) of the covariance matrix $\mathbf{\Sigma}$. The admissible set of k is $[\underline{k}, \bar{k}]$ and that of σ is $[\underline{\sigma}, \bar{\sigma}]$, which both are adaptive to the specific task. Details can be found in [10].

B. Task redundancies analysis

1) Learn null-space component ($\dot{\mathbf{k}}_{ns}$):

The human stiffness $\mathbf{k}$ have been estimated from Section IV-A. By differentiating $\mathbf{k}$, the observations now are $[\mathbf{k}, \dot{\mathbf{k}}]$ pairs. It is assumed that the impedance modulation contains task priorities ($\dot{\mathbf{k}} = \dot{\mathbf{k}}_{ts} + \dot{\mathbf{k}}_{ns}$). The goal of this section is to estimate $\mathbf{A}$ and learning $\dot{\mathbf{k}}_{ts}$. In order to learn constraint $\hat{\mathbf{A}}$, the task space component $\dot{\mathbf{k}}_{ts}$ needs to be filtered and an the null-space component $\dot{\mathbf{k}}_{ns}$ needs to be estimated. $\dot{\mathbf{k}}_{ns}$ is estimated by minimising the risk function

$$E[\dot{\mathbf{k}}_{ns}] = \sum_{n=1}^{N} ||\mathbf{P}_n \dot{\mathbf{k}}_n - \dot{\mathbf{k}}_{ns}||^2 \tag{7}$$

[1]Note that, using variance to estimate impedance may cause difficulties in close contact tasks [17]. In such situations, alternative approaches, such as looking at patterns of co-contraction of muscles through electromyography, may be used instead.

with $\mathbf{P}_n = \dot{\mathbf{k}}_n \dot{\mathbf{k}}_{ns}^T / ||\dot{\mathbf{k}}_{ns}||^2$. Here $\dot{\mathbf{k}}_n$ is the n-th data point from the data subset.

Minimising (7) is equal to minimising the difference between the current model of the null-space movement, and the observations *projected onto that model*. A detail illustration of the approach can be found in [4].

2) Learn constraint matrix $\mathbf{A}$:

After decomposing $\dot{\mathbf{k}}_{ts}$ from $\dot{\mathbf{k}}$, the task-space component can be readily derived by $\dot{\mathbf{k}}_{ts} = \dot{\mathbf{k}} - \dot{\mathbf{k}}_{ns}$, and the task constraint $\mathbf{A}$ can be learnt by minimising the inconsistency error [19]:

$$E[\mathbf{N}] = \sum_{n=1}^{N} ||\mathbf{k}_n - \mathbf{N}\mathbf{k}_n||^2 \tag{8}$$

Detail explanations can found in [19]. A state dependent extension for estimating $\mathbf{A}(\mathbf{k})$ can be found in [20].

C. Null-space controller optimisation

From §IV-B, the constraint is estimated as $\hat{\mathbf{A}}$ and the task critical components which decomposed from human demonstration are estimated as $\hat{\mathbf{u}}_{ts}$. In this section, the robot-specific controller is optimised in the null-space of the task-critical part of the human demonstration (as shown in (9)). To formulate the optimisation problem, it is essential to define a new state variable $\mathbf{v} = [\mathbf{x}, \mathbf{u}]^T$, where $\mathbf{u} = [\mathbf{k}, \mathbf{d}]^T$. A desired end-effector space dynamics (*e.g.,* point mass system) which is modulated by robot impedance controller can be rewritten in its constrained form

$$\dot{\mathbf{u}}_{ts} = \mathbf{A}^{\dagger}\mathbf{b} \tag{9}$$

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{g}(\mathbf{x}, \mathbf{u}) \\ \mathbf{A}^{\dagger}\mathbf{b} + \mathbf{N}\dot{\mathbf{u}}_0 \end{bmatrix} \tag{10}$$

where $\mathbf{A} = \mathbf{I} \otimes \hat{\mathbf{A}} (\mathbf{I} \in R^{\mathcal{P}/2 \times \mathcal{P}/2})$ and $\mathbf{N}$ are the augmented constrained impedance modulations since the impedance variables are uncoupled. $\dot{\mathbf{u}}_0$ is the null-space control command. Therefore, this problem is transformed as follow:

$$\begin{aligned} &\dot{\mathbf{v}} = h(\mathbf{v}, \dot{\mathbf{u}}_0) \\ s.t. \quad &J = l(\mathbf{v}(T)) + \int_{t=1}^{T} j(\mathbf{v}(t), \dot{\mathbf{u}}_0(t)) \end{aligned} \tag{11}$$

where $l(\mathbf{v}(T))$ is the terminal cost and $j(\mathbf{v}(t), \dot{\mathbf{u}}_0(t))$ is the running cost. Solving (11) using Iterative Linear Quadratic Regulator (ILQR) is essentially finding an optimal $\dot{\mathbf{u}}_0 = \boldsymbol{\pi}(\mathbf{u})$ that minimises J and produces no effect on primary task critical component ($\dot{\mathbf{u}}_{ts}$) learned from human demonstrations.

The control policy ($\boldsymbol{\pi}$) derived so far is in the operational space. In order to reproduce it on robot, it needs to be transformed to the joint space of the robot. In task space, the restoring force is:

$$\mathbf{f}_s = -[\mathbf{K}_s, \mathbf{D}_s]\delta\mathbf{z} \in \mathbb{R}^{\mathcal{S}} \tag{12}$$

where $\delta\mathbf{z} = \mathbf{J}(\mathbf{q})\delta\mathbf{q}$, $\mathbf{q} \in \mathbb{R}^Q$ is the the joint space variable. In joint space, a sequence of equilibrium joint-space position ($\hat{\mathbf{q}}$) can be resolved by:

$$\begin{aligned} \hat{\mathbf{q}}_{t+1} &= \hat{\mathbf{q}}_t + \delta t \delta\mathbf{q} \\ \delta\mathbf{q} &= \mathbf{J}(\mathbf{q})^\dagger \delta\mathbf{z} \end{aligned} \tag{13}$$

The transformation from task space stiffness and damping $\mathbf{K}_s$, $\mathbf{D}_s$ to joint space stiffness and damping $\mathbf{K}_q$, $\mathbf{D}_q$ is:

$$\begin{aligned} \mathbf{K}_q &= \mathbf{J}(\mathbf{q})^T \mathbf{K}_s \mathbf{J}(\mathbf{q}) \\ \mathbf{D}_q &= \mathbf{J}(\mathbf{q})^T \mathbf{D}_s \mathbf{J}(\mathbf{q}) \end{aligned} \tag{14}$$

The joint space robot impedance controller is:

$$\begin{aligned} \boldsymbol{\tau} &= \mathbf{J}(\mathbf{q})^T \mathbf{f}_q \\ \mathbf{f}_q &= -[\mathbf{K}_q, \mathbf{D}_q]\delta\mathbf{q} \in \mathbb{R}^Q \end{aligned} \tag{15}$$

Where $\boldsymbol{\tau}$, $\boldsymbol{\tau}_{ff}$ and $\delta\mathbf{q}$ are joint torque, feed-forward torque, and deviation of the current joint configuration from the reference $\hat{\mathbf{q}}$.

V. Result

In this section, the proposed method for teaching impedance behaviour by combining human and robot 'best practice' is implemented in both simulations and the real world.

A. Evaluations on simulated point mass dynamics

Experiment procedure: This is a mock up simulation where the human impedance $\mathbf{u}_i$ is generated. In the operational space, the dynamics of the hand end-point is assumed a 2-dimensional point mass (PM) model. To imitate the coupling effect of the stiffness and damping, the relation $\mathbf{u} = [k_1, k_2, 0.5k_1, 0.5k_2]$ is defined by coupling function $\mathbf{C}(\mathbf{k})$ with $d = 0.5k$. The primary task (J_1)[2] is a y-axis reaching task where the ground truth task constraint $\mathbf{A}$ and null-space control policy $\boldsymbol{\pi} = -2\mathbf{u}$ are assumed unknown. This null-space policy essentially tracks zero stiffness. It imitates that the person is getting less care about his stiffness in the nulls-space regardless the initial quantities. The problem is formulated as follow:

$$\begin{aligned} \dot{\mathbf{z}} &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{z} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1/m & 0 \\ 0 & 1/m \end{bmatrix} \mathbf{f} \\ \dot{\mathbf{k}} &= \mathbf{A}^\dagger \dot{\mathbf{b}} + \mathbf{N}\dot{\mathbf{k}}_0 \\ \mathbf{f} &= -\begin{bmatrix} \mathbf{x}_1 - \hat{\mathbf{x}}_1 & 0 & \dot{\mathbf{x}}_1 - \dot{\hat{\mathbf{x}}}_1 & 0 \\ 0 & \mathbf{x}_2 - \hat{\mathbf{x}}_2 & 0 & \dot{\mathbf{x}}_2 - \dot{\hat{\mathbf{x}}}_2 \end{bmatrix} \begin{bmatrix} \mathbf{k} \\ \mathbf{C}(\mathbf{k}) \end{bmatrix} \end{aligned} \tag{16}$$

[2]This objective function is typically used for evaluating human reaching performance [21].

$$\begin{aligned} s.t. \quad J_1 &= (\mathbf{z}_T - \hat{\mathbf{z}})^T \mathbf{Q} (\mathbf{z}_T - \hat{\mathbf{z}}) \\ &+ \int_0^T (\mathbf{z}_T - \hat{\mathbf{z}})^T \mathbf{Q} (\mathbf{z}_T - \hat{\mathbf{z}}) + \dot{\mathbf{b}}(t)^T \mathbf{R} \dot{\mathbf{b}}(t) \delta t \end{aligned} \tag{17}$$

where the parameters can be found in TABLE.I. The experiment is repeated 10 times. For each experiment, 10 demonstrations are simulated by running Iterative Linear Quadratic Regulator from a different initial state $\mathbf{v}_0$. After generation of the human demonstrations, the stiffness variables $[k_1, k_2]^T$ are decomposed into task-space components and null-space components. The training samples are state-action pairs $[\mathbf{k}, \dot{\mathbf{k}}]$. The null-space component ($\hat{\mathbf{k}}_{ns}$) is first learned so that the task constraint ($\hat{\mathbf{A}}$) can be learned afterward. The normalised unconstrained policy error (NUPE) and normalised projection error (NPE) are used to evaluate the learning performance of the $\hat{\mathbf{u}}_{ns}$. The normalised projected policy error (NPPE) and normalised projected observation error (NPOE) are used to evaluate the learning performance of $\hat{\mathbf{A}}$.

parameters		values
$\dot{\mathbf{b}} = \dot{k}2[N/m]$		$U(-50, 50)$
$\mathbf{v}_0$	$x_1[m]$	$U(-2, 2)$
	$x_2[m]$	$U(0, 2)$
	$\dot{x}_1[m/s]$	0
	$\dot{x}_2[m/s]$	0
	$[k_1, k_2, d_1, d_2]^T$	$U(1, 2)$
$\mathbf{z}_T$		$[1, 1, 0, 0]^T$
$\mathbf{Q}$		$\mathrm{diag}([0, 1000, 0, 0])$
$\mathbf{R}$		0.01
m[kg]		1.59
$\mathbf{A}$		$[0, 1]$

TABLE I: Parameters setups for the y-axis reaching task in simulation.

In the last step, the blended null-space controller $\boldsymbol{\pi}$ is learned by given the task-space component ($\mathbf{k}_{ts}$). The coupling function $\mathbf{u} = \mathbf{C}(\mathbf{k})$ takes no effect. The task component $\mathbf{u}_{ts} \equiv [k_1, 0.5k_1]^T = \mathbf{A}^\dagger \dot{\mathbf{b}}$. The secondary task objective J_2 is designed to stabilise in the x-axis and minimising the energy:

$$\begin{aligned} J_2 &= (\mathbf{z}_T - \hat{\mathbf{z}})^T \mathbf{Q} (\mathbf{z}_T - \hat{\mathbf{z}}) \\ &+ \int_0^T (\mathbf{z}_T - \hat{\mathbf{z}})^T \mathbf{Q} (\mathbf{z}_T - \hat{\mathbf{z}}) + \dot{\mathbf{u}}_0(t)^T \mathbf{R} \dot{\mathbf{u}}_0(t) \delta t \end{aligned} \tag{18}$$

where $\mathbf{Q} = \mathrm{diag}([0, 0, 1, 0])$ and $\mathbf{R} = \mathrm{diag}([0.01, 0.01, 0.01, 0.01])$, meaning the x-axis stability is of most concerned. 10 optimal trajectories $\dot{\mathbf{u}}_0$ are solved by using ILQR. A further simulation is made where a perturbation with force ($f \in [10N, -50N]^T$) is given at the steady state to illustrate the performance of the blended controller.

Result: The simulated demonstrations for 1 experiment are shown in Fig.3-a where all the end-point positions have reached the line target. In the bottom figure, the stiffness k_1 and damping d_1 profiles (solid line) for human impedance control are proportionally dependent (*i.e.,* the magnitude of damping is half).

For task redundancy analysis, the learning performance of $\hat{\mathbf{A}}$ and $\hat{\mathbf{u}}_{ns}$ are evaluated. NUPE and NPPE are the preferable metrics when all the ground truth information are available. NPE and NPOE are assumed the true $\mathbf{u}_{ns}$ and

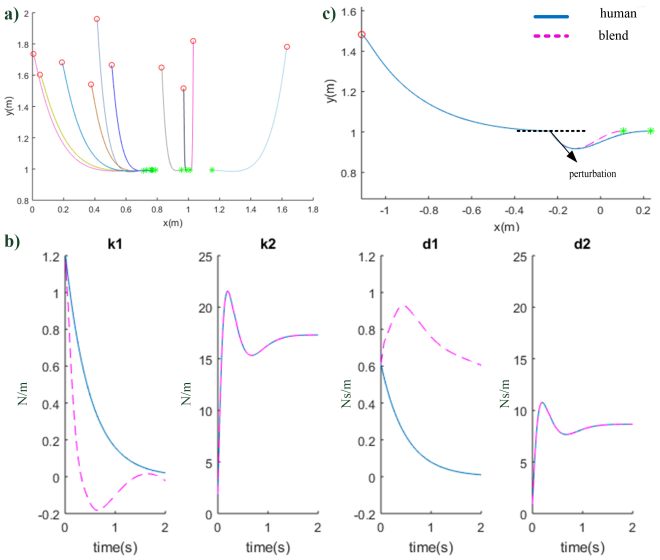

Fig. 3: a)Simulated end-point position sunder coupled dynamics. b) One trial of human impedance control and its corresponding blended controller. c) Simulated perturbation at the steady state using human and blended impedance controller (the perturbation is annotated by arrows).

Learning target	$\hat{\mathbf{u}}_{ns}$		$\hat{\mathbf{A}}$	
Metric	NUPE	NPE	NPPE	NPOE
Result	0.2 ± 0.002	1.1×10^{-4}	0	0

TABLE II: Evaluation of learning $\mathbf{u}_{ns}$ and $\mathbf{A}$. Results are mean and standard deviation ($mean \pm s.d.$) over 10 trials (the s.d. is negligible if no value appears in the result).

the unconstrained policy $\boldsymbol{\pi}$ are unknown. The results are summarised in TABLE.II. The low error indicates a good approximation of $\hat{\mathbf{A}}$ and $\hat{\mathbf{u}}_{ns}$.

For blended controller, an example of optimal trajectory is shown in the dot line in Fig.3-b. By comparing with human impedance profile, the blended controller shows the decouple behaviour of the stiffness and damping in the x-axis. The damping (d_1) is increasing and maintaining high compared with the human damping in solid line. The ($mean \pm s.d.$) are used to evaluate the performance of the human and blended controller according to J_2. They all have the same task-component but different null-space behaviour. The scores are 2.6 ± 1.7 and $0.02 \pm 8.9 \times 10^{-5}$ for human and blended controller respectively.

Discussion: The primary task objective is well accomplished in the simulated human controller. Using the demonstrations, the redundancies analysis successfully separates the task-space components and learns the task constraint. By imitating the $\mathbf{u}_{ts}$ and blending robot-specific controller, the blended controller outperforms the human controller in accomplishing secondary task objective. Furthermore, comparing with human, the blended controller generates stronger force against the perturbation. This also shows that the blended controller behave optimally and better than the human.

B. Evaluations on the Sawyer arm robot

Experiment procedure: The actual experiment is tested and evaluated on Sawyer robot[3]. The human operator is instrumented by a *Trakstar* sensor at the finger tip (top left in Fig.4). It generates positional information $\mathbf{x}_t$. 15 demonstrations are recorded. The number of state $c = 10$ was selected heuristically. For training the Gaussian mixture model, 10 trials are randomly shuffled and this process is repeated 10 times. For estimating stiffness, the $\bar{k}$ and $\underline{k}$ were experimentally selected as $200N/m$ and $50N/m$. The $\bar{\sigma}$ and $\underline{\sigma}$ are experimentally selected as 0.1 and 0.005. The damping is assumed to be coupled with the stiffness by relation $d = 0.5k$. For task redundancies analysis using demonstrated stiffness $\mathbf{k} = [k_1, k_2]^T$, the number of radial basis functions for both estimating $\hat{\mathbf{u}}_{ns}$ and $\hat{\mathbf{A}}$ are 10. For blending robot-specific controller, the same PM dynamics with constant m = 1.59 kg (approximate hand weight) is used. The equilibrium position value $\mathbf{x}_0$ is equal to $\bar{\mathbf{x}}$ which is the averaging path (produced by GMR) across all the samples because the stiffness profile is derived locally. For maximising the stability and minimising the energy, $\mathbf{Q} = \text{diag}([0, 0, 100, 0])$ and $\mathbf{R} = \text{diag}([0.01, 0.01, 0.01, 0.01])$ are heuristically chosen for task objective J_2. The admissible control input is $\dot{\mathbf{u}} \sim U(-100, 100)$. At steady state, 10 manual perturbations, $\tilde{\mathbf{f}} \in [-20, 20]$ (top middle plot in Fig. 4) are generated on both direction of the x-axis to evaluate the performance of blended controller. In order to quantify and evaluate the stability performance, the averaged state changing rate ($\gamma = 1/N \sum_{i=1}^{N} \delta x_i$, where $i = 1, ..., N$) at each timestamps on x-axis after each perturbation is recorded (the time interval is the response time).

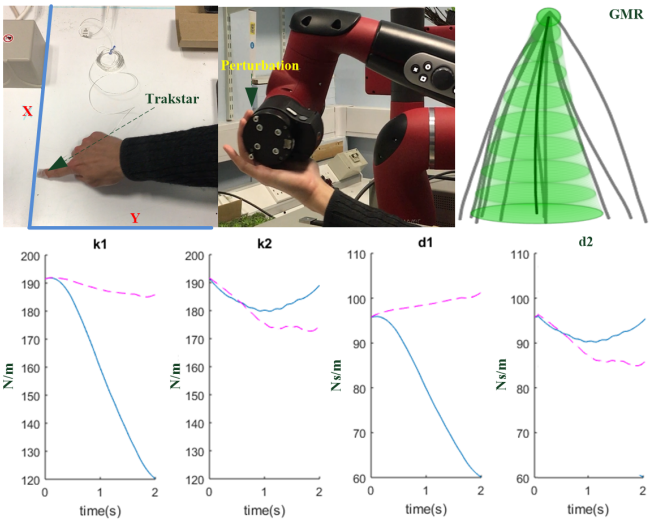

Fig. 4: Human is demonstrating the reaching task using tracking sensor (top left). A manual perturbation is given at steady state (top middle). Human demonstrations with Gaussian mixture model and GMR encoding (top right). An example of Human and blended impedance control(bottom).

Result: The covariance matrix for each time step is calculated by using GMR. As an example, in the top right of Fig.4, the GMR result shows that at the early stage, the covariance of the data is almost a circle, but gradually the x-axis of the ellipse is elongated. Since the stiffness is coupled with damping and it is inverse proportional to the covariance, one example of the impedance profile is shown in bottom of Fig.4 in solid line.

For task redundancy analysis, the NPE for learning $\hat{\mathbf{u}}_{ts}$ is 0.027, and the NPOE for learning $\hat{\mathbf{A}}$ is 0.036. The learned

[3]Details of the robot make and model can be found in `www.rethinkrobotics.com/sawfffffyer/`

	J_1	J_2
Human	$0.065 \pm 2.595 \times 10^{-6}$	$8.178 \times 10^{-4} \pm 1.765 \times 10^{-8}$
Blend	$0.064 \pm 2.421 \times 10^{-6}$	$7.450 \times 10^{-5} \pm 6.841 \times 10^{-11}$

TABLE III: Evaluation on task performance using human and blended impedance controller. Results are $(mean \pm s.d.)$ over 10 trials.

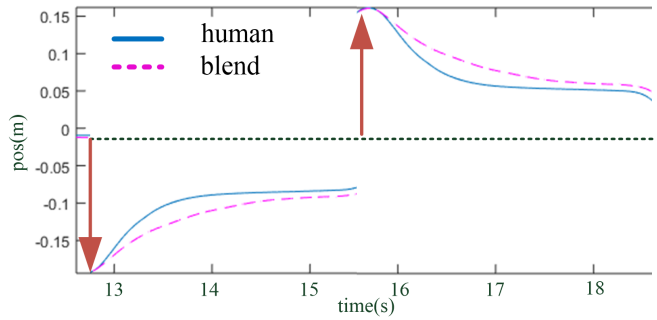

Fig. 5: An example of reproduced impedance profile using human demonstrations (solid) and blended controller (dash) under perturbation (arrow). The blended controller responses slower than human.

constraint $\hat{\mathbf{A}} = [0, 0.98]$. This is biased from the anticipated result $[0, 1]$.

By imitating the learned $\hat{\mathbf{u}}_{ts}$, one example of the blended controller profile is shown in bottom of Fig.4 in dash line. The task-space behaviour is still coupled but slightly different from the human. The null-space behaviour is decoupled and different from the human. A statistical analysis is shown in TABLE.III where the blended controller has accomplished the primary task with J_1 similar to human. On the other hand, the blended controller outperforms in the secondary task by comparing the values in J_2. In perturbation experiment, the mean and standard deviation values for using human impedance and blended controller are $0.09 \pm 0.338m/s$ and $0.0343 \pm 0.0068m/s$.

Discussion: The GMR plot indicates that the data contain much larger variation in the x-axis and the same y-axis target is reached which aligns with the primary task objective. However, the results in task redundancies analysis shows some bias in estimation. This is reasonable since the real world data is noisy and imperfect.The result in performance evaluations using J_1 and J_2 indicate that the primary task has been accomplished in both human and blended controller, but greater stability is achieved in blended controller for J_2. This behaviour have confirmed by giving manual perturbations in both statistical results in γ and visually in Fig. 5.

VI. Conclusion

In this paper, a programming by demonstration framework has been proposed for combining the human and robot 'best practice' for teaching impedance modulation task to robot. Both the simulation and experimental results show that the blended controller can accomplish the primary task and allow the stiffness and damping behave independently in the secondary task. The results (statistically and visually) show a clear improvement of the blended robot-specific controller. In future work, the proposed method will be applied to different manipulation tasks, *i.e.,* a hammering task where the task redundancy also appears. The learning method implemented in this paper does not require the user knows the task constraints nor the null-space components. However, it is possible the task constraint and the critical component is known beforehand, therefore, the user can directly separates the impedance demonstrations and implement the null-space optimisation directly. As a matter of fact, the proposed method is applicable to the in-contact manipulation task if the impedance profile can be measured (*i.e.,* through haptic or EMG device). This will be addressed in the future works.

References

[1] N. Hogan, "Impedance Control: An Approach to Manipulation," *American Control Conference*, 1985.

[2] R. F. Kirsch, D. Boskov, and W. Z. Rymer, "Muscle stiffness during transient and continuous movements of cat muscle: perturbation characteristics and physiological relevance," *IEEE TBME*, 1994.

[3] D. Braun, M. Howard, and S. Vijayakumar, "Exploiting variable stiffness in explosive movement tasks," in *Robotics: Sci. & Sys.*, 2011.

[4] C. Towell, M. Howard, and S. Vijayakumar, "Learning null space policies," in *Int. C. Robotics Res.*, 2010.

[5] G. Biggs and B. MacDonald, "A survey of robot programming systems," in *Proceedings of the Australasian Conference on Robotics and Automation*, 2003.

[6] S. Calinon, F. Guenter, and A. Billard, "On learning the statistical representation of a task and generalizing it to various contexts," in *ICRA*, 2006.

[7] S. Calinon and A. Billard, "Learning of gestures by imitation in a humanoid robot," Tech. Rep., 2007.

[8] A. J. Ijspeert, J. Nakanishi, and S. Schaal, "Learning attractor landscapes for learning motor primitives," in *NIPS*, 2003.

[9] F. Guenter, M. Hersch, S. Calinon, and A. Billard, "Reinforcement learning for imitating constrained reaching movements," *Adv. Robotics*, vol. 21, no. 13, pp. 1521–1544, 2007.

[10] K. Kronander and A. Billard, "Online learning of varying stiffness through physical human-robot interaction," in *ICRA*, 2012.

[11] S. M. Khansari-Zadeh and O. Khatib, "Learning potential functions from human demonstrations with encapsulated dynamic and compliant behaviors," *Auton. Robots*, vol. 41, no. 1, pp. 45–69, 2017.

[12] T. Petrič, L. Colasanto, A. Gams, A. Ude, and A. J. Ijspeert, "Bio-inspired learning and database expansion of compliant movement primitives," in *Humanoids*, 2015.

[13] T. Mori, M. Howard, and S. Vijayakumar, "Model-free apprenticeship learning for transfer of human impedance behaviour," in *Humanoids*, 2011.

[14] A. Ajoudani, N. G. Tsagarakis, and A. Bicchi, "Tele-impedance: Towards transferring human impedance regulation skills to robots," in *ICRA*, 2012.

[15] C. S. Rozo L, Bruno D and C. D. G, "Learning optimal controllers in human-robot cooperative transportation tasks with position and force constraints," in *IROS*, 2015.

[16] S. Haddadin, A. Albu-Schäffer, O. Eiberger, and G. Hirzinger, "New insights concerning intrinsic joint elasticity for safety," in *IROS*, 2010.

[17] P. L and P. T, "Human-in-the-loop approach for teaching robot assembly tasks using impedance control interface," in *ICRA*, 2015.

[18] S. Calinon, "A tutorial on task-parameterized movement learning and retrieval," *Intelligent Service Robotics*, vol. 9, no. 1, pp. 1–29, 2016.

[19] H.-C. Lin, M. Howard, and S. Vijayakumar, "Learning null space projections," in *ICRA*, 2015.

[20] H.-C. Lin, P. Ray, and M. Howard, "Learning task constraints in operational space formulation," in *ICRA*, 2017.

[21] E. Todorov and M. I. Jordan, "Optimal feedback control as a theory of motor coordination," *Nature neuroscience*, 2002.